# TEMPORAL EXEMPLAR CHANNELS IN HIGH-MULTIPATH ENVIRONMENTS

*Mohamed Kashef*[⋆] *Peter Vouras*[†] *Robert Jones*[‡] *Richard Candell*[⋆] *Kate A. Remley*[‡]

[⋆] Engineering Laboratory, National Institute of Standards and Technology (NIST), Gaithersburg, Maryland, United States

[†] Communications Technology Laboratory, NIST, Gaithersburg, Maryland, United States

[‡] Communications Technology Laboratory, NIST, Boulder, Colorado, United States

## ABSTRACT

Industrial wireless plays a crucial role in cyber-physical system (CPS) advances for the future vision of smart manufacturing. However, industrial wireless environments are different from each other and are different from home and office environments. Hence, industrial wireless channel modeling is essential for the development of industrial wireless systems. Moreover, millimeter-wave (mmWave) wireless bands have a high potential to be used for the high data-rates required for industrial automation reliability, with multiple antennas envisioned to mitigate the high path loss. As a result, in this work, we introduce a machine learning (ML) based exemplar extraction approach on mmWave wireless spatial-channel measurements. The proposed approach processes the measured power-angle-delay-profiles to cluster them into a number of groups with respect to the angle of arrival. Then, an exemplar power-delay-profile (PDP) is extracted to represent each group. The resulting set of exemplars provide a tractable way to conduct mmWave industrial wireless systems testing and evaluation by compactly representing various feature groups. This allows the assessment of wireless equipment over the exemplars without the need to test over all of the different instances of wireless channel paths.

***Index Terms—*** Channel modeling, clustering, exemplar channel, industrial wireless, unsupervised learning

## 1. INTRODUCTION

In future industrial systems, wireless-communication technologies will play a critical role in achieving massive connectivity between various operational components and allowing easier equipment mobility. Industrial physical environments are different than office and home environments which leads to different wireless channel characteristics such as the achievable delay and reliability [1,2]. Generic models are being studied for indoor industrial channels such as [3] where four different categories of wireless channels in indoor factories are considered. However, various industrial environments differ from each other in their layouts, types of equipment, and the performed industrial activities. Hence, designing and testing of industrial wireless systems requires knowledge of the channel characteristics of the corresponding environment [4].

The limited availability of sub-6 GHz wireless spectrum motivated the utilization of millimeter-wave (mmWave) bands for many new wireless technologies. In July 2016, the United States Federal Communications Commission (FCC) allocated 3.85 GHz of licensed spectrum near 28, 37, and 39 GHz for 5G mobile networks, and 7 GHz of unlicensed spectrum from 64–71 GHz that is adjacent to the existing 57–64 GHz unlicensed Industrial, Scientific, and Medical (ISM) bands. The channel characteristics in the mmWave bands are different than sub-6 GHz bands. Moreover, with many licensed bands, they offer a potential candidate for industrial wireless. As a result, channel models are required based on measurements taken in various environments to study the behavior of several technologies and equipment communicating in the mmWave bands.

The power delay profile (PDP) of a wireless channel captures the temporal variations of the channel due to multipath components (MPCs) [5]. The classification and clustering of wireless channels, depending on the extracted temporal features from the corresponding PDPs using machine learning (ML) approaches, have been studied in the literature such as in [6–10]. In these papers, both supervised and unsupervised learning were used for scenario identification. To the best of our knowledge, we are the first to use a clustering ML approach to study the directional impact of the wireless channel and obtain exemplars that can be used to test devices and technologies while representing various groups of the measured directions. The contribution of the paper is introducing the approach, and results that emphasize benefits of applying the innovative approach over realistic channel measurements.

In this work, we introduce an unsupervised learning approach to obtain exemplar PDPs from mmWave channel measurements. Synthetic-aperture measurements are used to determine the channel's power-angle-delay profile (PADP), which characterizes both the angle-of-arrival and time-of-arrival of received power. Synthetic apertures typically consist of a single antenna element that is scanned over multiple locations and the complex fields are acquired over the lattice. The spatial characteristics of the channel are reconstructed in post processing by combining measurements and applying various phasing combinations, much like a phased array. Synthetic apertures have a number of benefits for characterizing spatial channels, including high dynamic range, lower reflectivity (due to the use of a single antenna element), and lower cost.

The proposed approach serves as a way to compactly represent various directional feature groups. Moreover, the result of the proposed approach allows the test and assessment of wireless equipment over the exemplars without the need to test over all of the different instances of wireless channel paths or to evaluate the performance over a generic model that does not capture the specifics of a certain environment. This approach can capture the environmental cases that may stress wireless device performance and potentially reveal flaws or deficiencies that can be improved upon in future device design iterations. To this end, we chose a very challenging propagation scenario inside the Central Utility Plant (CUP) that generates a lot of dense multipath reflections. Unsupervised ML is used to partition the measured PDPs into clusters that correspond to different directions and to extract canonical PDPs that embody all the salient features of each cluster. These exemplars can be used as device stressing design points that can be well replicated in millimeter wave test chambers or within simulation frameworks to test actual device performance. The exemplars generalize the characteristics of measured site-specific data without having to resort to a statistical averaging approach that ultimately yields only benign channel conditions but also requires extraordinary amounts of measured data. Another advantage of the proposed ML approach is that it enables the capability to emphasize or de-emphasize the environmental features that might be more or less important for testing particular device capabilities.

**Fig. 1**. The CUP measurement environment

## 2. PROBLEM DESCRIPTION

In this section, an overview of the data collection process is presented. The data preparation and the format of the resulting data are described. Finally, the addressed problem in this work is stated.

### 2.1. Environment and Data Collection

Measurements were performed in the highly reflective CUP at the Department of Commerce Boulder Laboratories in 2019. This environment consists of large boiler tanks, piping, and numerous racks of control hardware, as shown in Fig. 1. The vector network analyzer (VNA) was placed in a small rack located between the transmit horn antenna and the synthetic aperture receive array. The results presented in this paper were obtained from measurements over the WR-28 waveguide band of 26.5 GHz to 40 GHz with a dynamic range typically on the order of about 90 dB. We positioned a directional horn transmit antenna to point directly at a bank of switches and tanks, as shown in Fig. 2. The horn antenna provides 17 dBi gain, is linearly polarized, and has a 23°/24° 3 dB beamwidth in the E/H planes. It was oriented upward at an elevation angle of approximately 15°. Our synthetic aperture receive array was oriented toward the switch bank as well, minimizing the line-of-sight component. We configured the synthetic aperture [11] to scan a 35-by-35 planar grid with 3mm spacing between the sample points ($\lambda/2$ at 40 GHz). The minimum beamwidth of the synthetic aperture array response is 2.9° attained at 40 GHz in the boresight direction. Since the array is square, the azimuth and elevation beamwidths are equal. With this, we reconstructed directional PDPs corresponding to a beamwidth of 2.9° in azimuth and elevation.

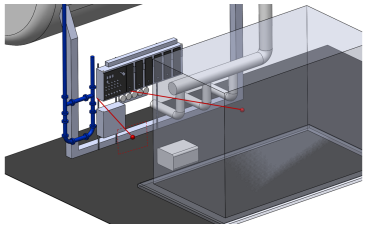


**Fig. 2**. CAD model of the CUP environment. The red lines show the boresight directions for the transmit antenna and synthetic aperture array. The transparent rectangular prism corresponds to the boiler that is opposite of the control panel. The dotted red rectangle shows the planar synthetic aperture in the environment. The white rectangular prism on the ground corresponds to the VNA.

### 2.2. Data Preparation and Resulting Data

The $S_{21}$ parameters are collected by the synthetic-aperture channel measurement system in 10 MHz frequency increments between 26.5 and 40 GHz at every spatial sample. The $S_{21}$ parameters are processed using true time delay beamforming to steer the array mainbeam as described in [11]. Initially, a low-sidelobe taper is applied across the aperture that is frequency invariant in the boresight direction. Then to steer the array mainbeam towards a desired direction, an additional phase taper is applied across the aperture that varies linearly with frequency. After coherently combining the product of measured $S_{21}$ values and complex beamforming weights across all the aperture spatial samples, an inverse Fourier transform is utilized to transform the frequency domain data to the temporal domain. The result is known as a directional PDP.

Often, it is desirable that the sum of the power patterns of all the individual beams yields omnidirectional gain. Hence, the pointing directions specified at the peak of the mainbeam are chosen systematically using the approach described in [12] such that all beams overlap at the 3-dB beamwidth. This algorithm accounts for the fact that the width of the array mainbeam increases in proportion to the product of the cosines of the azimuth and elevation scan angles.

### 2.3. Problem Statement

We denote the PADP instants by $h(\theta, \phi, \tau)$, where $\theta$ and $\phi$ denote the azimuth and elevation of the angle of arrival, respectively, and $\tau$ is the delay. The collected data are sampled versions of the PADP, where $\theta$ and $\phi$ take discrete values in the ranges $\theta_{\text{min}} \leq \theta \leq \theta_{\text{max}}$ and $\phi_{\text{min}} \leq \phi \leq \phi_{\text{max}}$. As a result, the input of our problem is a set, $\mathcal{H}$, of the PDPs $h(\theta, \phi, \tau)$ for all fixed combinations of $\theta$ and $\phi$. The output of the proposed approach is $N$ disjoint groups $\mathcal{H}_i, i \in \{1, N\}$ using unsupervised ML clustering based on a set of defined features. Each group is represented by an exemplar PDP denoted by $\hat{h}_i(\tau)$, which represents the corresponding group $\mathcal{H}_i$.

## 3. MACHINE LEARNING APPROACH

In this section, the PDP exemplar extraction approach is detailed including the feature selection process and the ML clustering scheme.

### 3.1. Feature Extraction

Feature-based clustering generally can work with a large set of the PDPs features while it is only performed over a small set of features, in this work, to examine the approach and allow easier visualization of the results. The PDP for a certain pair of $\theta$ and $\phi$ is defined as

$$h(\theta, \phi, \tau) = \sum_{l=0}^{L-1} \alpha_l \delta(\tau - \tau_l), \quad \theta_{\text{min}} \leq \theta \leq \theta_{\text{max}}, \phi_{\text{min}} \leq \phi \leq \phi_{\text{max}} \quad (1)$$

where $\alpha_l$ is the power gain for the $l$-th path, $\tau_l$ is the path arrival time, $L$ is the number of the arrival paths, and $\delta(\tau)$ is the Dirac function.

Now, we define the features of $h(\theta, \phi, \tau)$ that we used to characterize all the directional PDPs. All of the features are evaluated for a single PDP at a certain pair of $\theta$ and $\phi$. We will drop the argument to simplify the expressions. The channel gain, $G$, is evaluated as

$$G = \sum_{l=0}^{L-1} \alpha_l. \quad (2)$$

The dynamic range of the PDP is defined as the ratio of the largest MPC to the smallest MPC. It is evaluated in dB as

$$R = 10 \log \frac{\max_{0 \leq l \leq L-1}(\alpha_l)}{\min_{0 \leq l \leq L-1}(\alpha_l)}. \quad (3)$$

The mean delay is the first moment of the power delay profile and is evaluated as follows

$$\tau_{\text{mean}} = \frac{1}{G}\sum_{l=0}^{L-1}\alpha_l\tau_l. \tag{4}$$

The root-mean-squared (RMS) delay spread is the second moment of the power delay profile and is evaluated as follows

$$\tau_{\text{RMS}} = \sqrt{\frac{1}{G}\left(\sum_{l=0}^{L-1}\alpha_l\tau_l^2 - \tau_{\text{mean}}^2\right)}. \tag{5}$$

### 3.2. Data Clustering

The first step in clustering is to normalize the features to the range of 0 to 1. Then, we perform spectral clustering over the vectors of features corresponding to each beam direction. Generally, spectral clustering is a powerful technique in clustering nonlinear and inseparable data and where non-convex clusters are allowed [13]. We have used the Scikit-learn implementation for the spectral clustering algorithm [14]. We denote the feature vector, corresponding to a PDP in a specific beam direction defined by a pair of azimuth and elevation angles, by $Z_j$, where $j$ is the index of the PDP. The length of the vector $Z_j$ is the number of features, $N_{\text{f}}$, and the total number of PDPs to be clustered is $N_{\text{H}}$.

The spectral clustering algorithm requires a similarity metric and a number of output clusters. In this work, we consider linear similarity although other similarity metrics can be considered as well. The pairwise linear similarity between feature vectors evaluated through their dot product may be given as

$$SL_{jk} = Z_j^{\text{T}} Z_k, \tag{6}$$

where $(*)^{\text{T}}$ is the transpose of the vector.

Finally, the number of the clusters is obtained through a recursive search for the maximum Silhouette score [15]. The Silhouette score measures how closely-related an object is to its own cluster against the other clusters. We repeat clustering over various values of $N$ and keep the clusters that achieve the highest Silhouette score.

### 3.3. Exemplar Extraction

The last phase of the approach is to extract an exemplar PDP from each of the clusters. The exemplar is a member of a cluster of PDPs. The initial used scheme defines the exemplar as the PDP with shortest mean feature-distance to other cluster members. In a cluster $\mathcal{H}_i$, the exemplar PDP, denoted by $\hat{h}_i(\tau)$, is the one corresponding to the feature vector $\hat{Z}_i$. The value of $\hat{Z}_i$ is evaluated as

$$\hat{Z}_i = \underset{Z_k \text{ s.t. } k\in\mathcal{H}_i}{\arg\min} \frac{1}{|\mathcal{H}_i|-1} \sum_{\substack{Z_j\neq Z_k \\ Z_j \text{ s.t. } j\in\mathcal{H}_i}} \sqrt{||Z_k - Z_j||^2}. \tag{7}$$

This exemplar represents the closest cluster member to the center of mass of the cluster where the masses of all the clusters members are equal.

Moreover, various weighting factors can impact the exemplar extraction process, where the mass of each member of the cluster can vary based on the importance of various features. One example is to scale the cluster members' mass based on their total channel power gain value, $G$. Hence, we introduce, to the exemplar extraction process, a channel-gain based scaling with a power-exponent tuning. We define power-exponent $\sigma$ to be the exponent to the PDP power gain $G$ that is multiplied by the minimization argument in (7). The extracted exemplar in this case with a power exponent of $\sigma$ is evaluated as follows

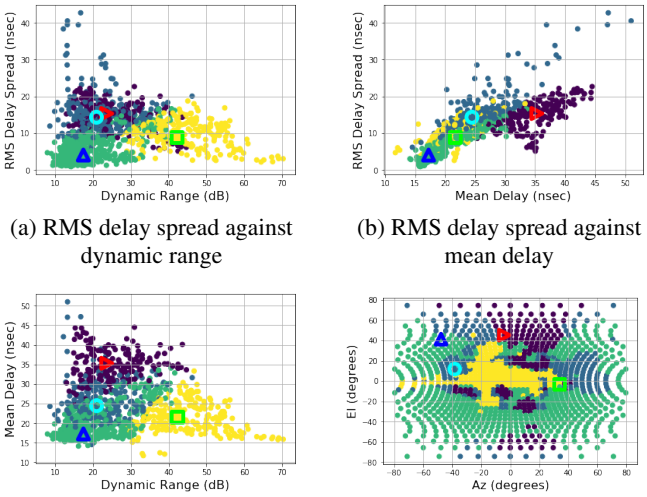
$$\hat{Z}_i = \underset{Z_k \text{ s.t. } k\in\mathcal{H}_i}{\arg\min} \frac{G_k^{\sigma}}{|\mathcal{H}_i|-1} \sum_{\substack{Z_j\neq Z_k \\ Z_j \text{ s.t. } j\in\mathcal{H}_i}} \sqrt{||Z_k - Z_j||^2}, \tag{8}$$

where $G_k$ is the total channel-power gain of the PDP corresponding to $Z_k$. For high power-exponent values, this scheme will pick the exemplar with the highest channel-power-gain value in the cluster.

## 4. RESULTS

In this section, we show the results obtained by running the exemplar extraction approach over the channel measurements. Two sets of results are included: i) Linear similarity metric with $\sigma = 0$, and ii) Linear similarity metric with $\sigma = 0.5$. In each set of results, we show two figures. The first shows the features of the resulting clusters through four scatter plot sub-figures: a) RMS delay spread (ns) against dynamic range (dB), b) RMS delay spread (ns) against mean delay (ns), c) mean delay (ns) against dynamic range (dB), and d) the directions in azimuth and elevation (degrees) of clusters. All cluster members have the same color with a corresponding exemplar marked with a specific shape. In the second figure, the PDPs of the exemplars are shown.

### 4.1. Linear Similarity with $\sigma = 0$




(a) RMS delay spread against dynamic range

(b) RMS delay spread against mean delay

(c) Mean delay against dynamic range

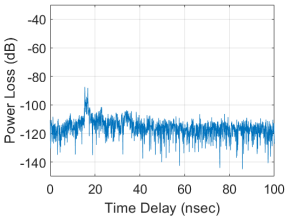
(d) Azimuth against elevation angles

**Fig. 3**. Clustered data with marked exemplars at $\sigma = 0$

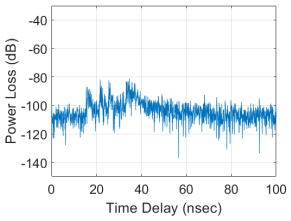
In Fig. 3, we draw the clustering result with the exemplars marked on the figures. In this case, we found that the optimal number of clusters is four, where the Silhouette score is 0.32. The value of Silhouette score can be used to demonstrate the effectiveness of the proposed clustering approach, because it measures how closely related a directional PDP is to its own cluster against the other clusters with respect to the used features.

The first cluster, in green, is characterized by low mean delay, low RMS delay spread, and low dynamic range. The second cluster, in purple, is characterized by high mean delay, low RMS delay spread, and low dynamic range. The third cluster, in blue, is characterized by high RMS delay spread, and low dynamic range.

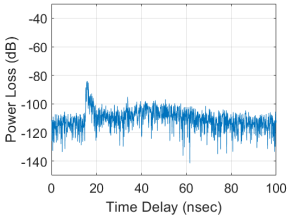
The fourth cluster, in yellow, is characterized by low mean delay, low RMS delay spread, and high dynamic range. The corresponding exemplars are marked in Fig. 3 with different shapes, namely, up-pointing triangle, right-pointing triangle, circle, and square, respectively. The PDPs of these exemplars are shown in Fig. 4. This figure shows that an exemplar PDP reflects the characteristics of the corresponding cluster. As an example, in Fig. 4(a), the PDP reflects the case where the PDP has a single peak, which leads to low mean delay and low RMS delay spread. This peak also has a low channel gain such that the dynamic range is low as well.



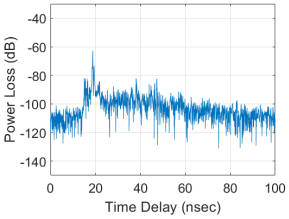

(a) Marked with △




(b) Marked with ▷



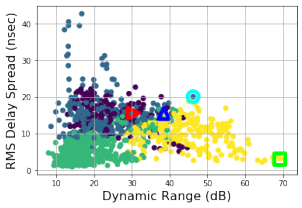

(c) Marked with ◯



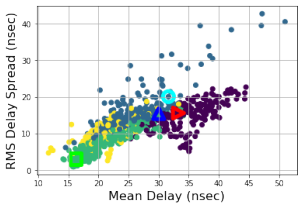

(d) Marked with □

**Fig. 4**. Marked exemplar PDPs corresponding to Fig. 3

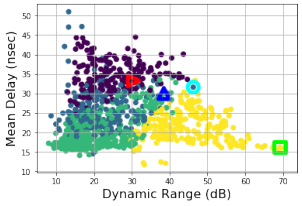
### 4.2. Linear Similarity with $\sigma = 0.5$



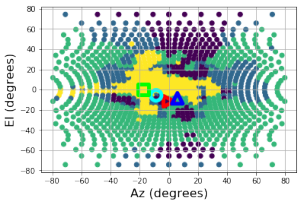

(a) RMS delay spread against dynamic range




(b) RMS delay spread against mean delay



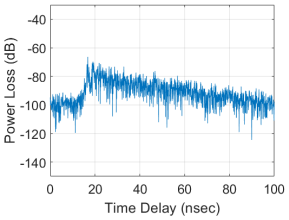

(c) Mean delay against dynamic range



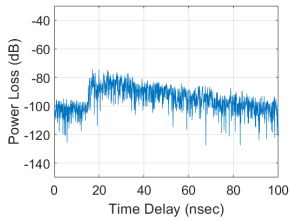

(d) Azimuth against elevation angles

**Fig. 5**. Clustered data with marked exemplars at $\sigma = 0.5$

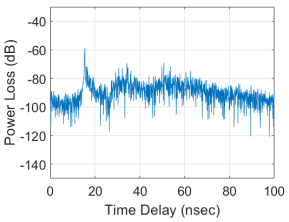
In this set of results, we present in Fig. 5 similar clusters as those obtained in Fig. 3. However, the increase in the exemplar extraction power exponent allows us to extract exemplars with higher channel-power gain and, hence, higher dynamic range. The clusters for the case of $\sigma = 0.5$ are shown in Fig. 5 and the corresponding exemplars are shown in Fig. 6, where they exhibit higher gain in the MPCs than the MPCs in Fig. 4. Note that the exemplars in Fig. 6 (a) and (b) look similar because the change in power-exponent moved the exemplars to the edges of the clusters and hence close to each other.



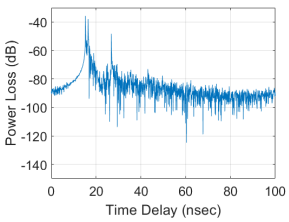

(a) Marked with △




(b) Marked with ▷




(c) Marked with ◯




(d) Marked with □

**Fig. 6**. Marked exemplar PDPs corresponding to Fig. 5

### 4.3. Comparison to the Average PDP

In Fig. 7, we show the average PDP of all the measured PDPs as a benchmark for representing the measured data. Comparing the exemplars to the average PDP, we observe that each exemplar represents a group of PDPs with certain features and hence can be used for testing wireless devices under certain conditions.

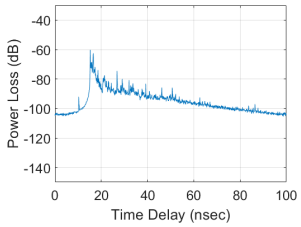



**Fig. 7**. The average PDP over all the directional PDPs

## 5. CONCLUSIONS

In this paper, we introduced an approach for directional PDP exemplar extraction from measured data for a static, highly reflective channel. The approach deploys unsupervised spectral clustering for PDP clustering and uses the delay and power characteristics for exemplar extraction. We have demonstrated that the wireless channel paths between two points in an industrial environment can have different characteristics depending on orientation. Hence, to operate in such an environment, a wireless node has to be tested under all different types of channel characteristics. The proposed approach serves as a way to compactly represent various feature groups. This allows the test and assessment of wireless equipment over the exemplars without the need to test over all of the different instances of wireless channel paths or to evaluate the performance over a generic model that does not capture the specifics of a certain environment.

**Disclaimer** Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

## 6. REFERENCES


[1] Kay Soon Low, Nu Win, and Meng Joo Er, "Wireless Sensor Networks for Industrial Environments," 2005.

[2] M. Cheffena, "Propagation channel characteristics of industrial wireless sensor networks [wireless corner]," *IEEE Antennas and Propagation Magazine*, vol. 58, no. 1, pp. 66–73, Feb. 2016.

[3] 3GPP, "Addition of indoor industrial channel model to Doc. 38.901 - Study on channel model for frequencies from 0.5 to 100 GHz," Tech. Rep. R1-1909807, 3rd Generation Partnership Project (3GPP), 09 2019.

[4] R. Candell, M. Kashef, Y. Liu, K. B. Lee, and S. Foufou, "Industrial wireless systems guidelines: Practical considerations and deployment life cycle," *IEEE Industrial Electronics Magazine*, vol. 12, no. 4, pp. 6–17, Dec. 2018.

[5] B. Soret, M. C. Aguayo-torres, and J. T. Entrambasaguas, "Capacity with explicit delay guarantees for generic sources over correlated Rayleigh channel," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 1901–1911, June 2010.

[6] J. Zhang, L. Liu, Y. Fan, L. Zhuang, T. Zhou, and Z. Piao, "Wireless channel propagation scenarios identification: A perspective of machine learning," *IEEE Access*, vol. 8, pp. 47797–47806, 2020.

[7] M. I. AlHajri, N. T. Ali, and R. M. Shubair, "Classification of indoor environments for iot applications: A machine learning approach," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 12, pp. 2164–2168, 2018.

[8] C. Huang, A. F. Molisch, R. Wang, P. Tang, R. He, and Z. Zhong, "Angular information-based NLOS/LOS identification for vehicle to vehicle MIMO system," in *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2019, pp. 1–6.

[9] Haihan Li, Yunzhou Li, Shidong Zhou, and Jing Wang, "Wireless channel feature extraction via GMM and CNN in the tomographic channel model," *Journal of Communications and Information Networks*, vol. 2, no. 1, pp. 41–51, Mar. 2017.

[10] Yunlong Yu, Fuxian Liu, and Sheng Mao, "Fingerprint extraction and classification of wireless channels based on deep convolutional neural networks," *Neural Processing Letters*, vol. 48, no. 3, pp. 1767–1775, Feb. 2018.

[11] P. Vouras, B. Jamroz, J.T. Quimby, A. Weiss, D.F. Williams, R. Leonhardt, D. Guven, R.D. Jones, J. Kast, and K.A. Remley, "Wideband synthetic aperture millimeter-wave spatial channel measurements with uncertainties," *in review*, 2020.

[12] G. V. Trunk W. M. Waters, D. P. Patel, "Optimum number of faces of a volume-scanning active array radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 3, July 1998.

[13] Soheila Ashkezari Toussi and Hadi Sadoghi Yazdi, "Feature selection in spectral clustering," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 4, no. 3, pp. 179–194, Sep 2011.

[14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[15] "Online, sklearn.metrics.silhouette_score, URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html," (visited: 2020-09-23).